Application Number	Application/Control No.	Applicant(s)/Patent Under Reexamination
	10/808,326	SHEN, SHIOUPYN
	Examiner	**Art Unit**
	Garrett A. Smith	2168

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/808,326	03/25/2004	Shioupyn Shen	0026-0072	5423

44989 7590 02/08/2008

HARRITY SNYDER, LLP
11350 Random Hills Road
SUITE 600
FAIRFAX, VA 22030

EXAMINER	
SMITH, GARRETT A	
ART UNIT	**PAPER NUMBER**
2168	

MAIL DATE	DELIVERY MODE
02/08/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/808,326	SHEN, SHIOUPYN	
	Examiner	Art Unit	
	Garrett A. Smith	2168	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *21 November 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-34* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-34* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This Office action is regarding Applicant's response filed 21 November 2007 to a prior Office action. Claims 1 – 34 are pending. Claims 1, 2, 5 – 11, 13 and 20 – 34 are amended.

2. This Office Action is the Second Action, Non-Final Rejection.

Response to Arguments

35 USC § 101

3. Applicant's arguments (page 11, fourth paragraph) and amendments, filed 21 November 2007, regarding the rejection under 35 USC § 101 of claim 20 – 30 and 34 have been fully considered but they are not persuasive. The recitation of a "computer-implemented device" can be reasonably interpreted to include software. Each of the recited elements of the claims can be implemented wholly in software. The Examiner notes, however, that recitation in the claims of a component which can only be reasonably interpreted as hardware may result in the claims being deemed statutory. Software, by itself, is not statutory. For these reasons, the rejection under 35 USC § 101 of claim 20 – 30 and 34 is **maintained**.

4. Applicant's arguments (page 11, fourth paragraph) and amendments, filed 21 November 2007, regarding the rejection under 35 USC § 101 of claims 31 – 33 have been fully considered and are persuasive. For these reasons, the rejection under 35 USC § 101 of claims 31 – 33 is **withdrawn**.

35 USC § 102(a): Scheimer et al

5. Applicant's arguments (pages 12 – 16) and amendments, filed 21 November 2007, regarding the rejection under 35 USC § 102(a) of claims 1, 2, 5, 8, 11, 13, 16, 17, 20, 24, 26, 27 and 28 have been fully considered but they are not persuasive.

Applicant argues (on page 12, first full paragraph), "SCHLEIMER does not disc lose or suggest choosing a subset of the plurality of overlapping blocks, as recited in claim 1." The Examiner respectfully disagrees. First, the Examiner notes that "overlapping" is not explicitly defined. There is only the mention of "overlapping" in the summary of the Specification and the Specification is silent as to what Applicant intends "overlapping" of blocks to entail. Therefore, a broadest reasonable interpretation is taken. Second, Applicant states, "SCHLEIMER appears to hash all of the 5-grams derived from the text and does not choose a subset of the 5-grams". The Examiner agrees with this assessment of Scheimer et al. The Examiner would further point to figure 2(b) and note that the 5-grams include overlapped section of text. Therefore when the hashing occurs in figure 2(c) an "overlapped" hash is created. The "overlapped" hash is grouped (and those groups are overlapping as well.

Applicant argues (on page 13, first and second full paragraphs), "Applicant submits that windows of hashes depicted in Fig. 2(e) of SCHLEIMER are not obtained by sampling the document." First, the Examiner would like to note that "to obtain a plurality of overlapping blocks" of claim 1 is the intended use of "sampling the document" as such is given little or no patentable weight. Second, the entire winnowing process is a sampling process as the hashes are based on the document itself.

"Sampling" is not defined nor claimed as a specific type of sampling is claimed. Therefore a broadest reasonable interpretation is taken. Sampling is defined as "the act or process of selecting a sample for testing, analyzing, etc" by Dictionary.com Unabridged (v 1.1). As such, the winnowing process is a sampling process and as such meets the limitation.

Applicant argues (on page 13, last paragraph), "Claim 2 recites that compacting the subset of the plurality of overlapping blocks includes setting bits in the representation of the document based on the subset of the plurality of overlapping blocks. SCHLEIMER does not disclose or suggest this feature." The Examiner respectfully disagrees. From figure 2(e), it is noted at this point in the process there exists windows of hashes (in the case of the example 14 windows). At figure 2(f), there is a "compaction" from the representation windows of hashes to a set of (in this example) 5 hashes. The act of placing (or even selecting) this hashes involves storage in memory which in turn involves the setting of bits. Once these hashes are placed together a representation of the document is formed in memory. Therefore, this limitation is taught.

For these reasons, the rejection under 35 USC § 102(a) of claims 1, 2, 5, 8, 11, 13, 16, 17, 20, 24, 26, 27 and 28 is **maintained**.

35 USC § 102(b): Burrows

6. Applicant's arguments (page 16 – 18) and amendments, filed 21 November 2007, regarding the rejection under 35 USC § 102(b) of claim 20 have been fully considered but they are not persuasive.

Applicant argues that Burrows does not disclose the elements of claim 20. The Examiner points to the Summary of the Invention of Burrows, specifically col 1, lines 48 – 54: "The invention provides a computer implemented method for indexing duplicate information stored as records having different unique addresses in a database. **The method generates a fingerprint for each record**. The fingerprint is a singular value derived from all of the information of the record according to a predetermined combination of the information of the record." As such, a processor unit in the distributed computer system 100 (see figure 1) executes these steps the processor is the fingerprint creation module. Compaction (not creation) is accomplished as shown in Figure 5.

For these reasons, the rejection under 35 USC § 102(b) of claim 20 is **maintained**.

35 USC § 103(a): Burrows in view of Ward et al

7. Applicant's arguments (pages 18 – 24) and amendments, filed 21 November 2007, regarding the rejection under 35 USC § 103(a) of claims 1, 2, 5 – 8, 13, 16, 17, 24 – 29, 31, 32 and 34 have been fully considered but they are not persuasive.

Applicant argues (on page 19, first full paragraph) that Burrows does not provide for a selection of a subset of overlapping blocks. As a "subset" can be of any number

elements in a "set", this also includes when all the elements are selected. There is no requirement in the claims that the number of elements selected be less then the number of elements in the set. Therefore, Burrows teaches this limitation.

Applicant argues (on page 19, second and third full paragraphs) that neither Burrows nor Ward et al provides for compaction. First, the Examiner would like to note that "to obtain a representation of the document" of claim 1 is the intended use of "compacting the subset of the plurality of overlapping blocks" as such is given little or no patentable weight. Second, col 9, lines 33 – 40 of Burrows states, "The parsing module 30 produces a sequence of pairs 500 in a collating order according to the location of the words 300 of the various pages 200." Also col 9, lines 55 – 67, discusses figure 6 in which a compressed data structure 71 is shown to describe the words parsed from the document. As such, Burrows describes this limitation.

The Examiner notes that Applicant submits that the Examiner's motivation to combine Burrows and Ward et al is insufficient for establishing a *prima facie* case of obviousness. However, the Examiner submits that each of the elements of claim 1 is well known as shown by Burrows. Ward shows a well known technique for using a windowing system for overlapping data in a fingerprint generation method/device. As such, a person of ordinary skill in the art would have recognized the use of this windowing method with that of Burrows because it yields the predictable result of quicker indexing and a higher accuracy of the resulting samples. See *KSR International Co. v. Teleflex Inc., 550 U.S.--, 82 USPQ2d 1386 (2007).*

For these reasons, the rejection under 35 USC § 103(a) of claims 1, 2, 5 – 8, 13, 16, 17, 24 – 29, 31, 32 and 34 is **maintained**.

Claim Rejections - 35 USC § 101

8. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

9. Claims **20 – 30 and 34** are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

10. In regard to **claim 20**, the Specification (¶48) provides evidence that Applicant intends the "fingerprint creation module" and the "similarity detection component" to be implemented in software. The "search engine" of dependant claim 21 is also intended to be implemented in software.

11. In regard to **claim 26**, the Specification (¶48) provides evidence that Applicant intends the "means for sampling ...", "means for choosing ..." and "means for compacting ..." to be implemented in software. The "means for calculating ..." of dependant claim 27 is also intended to be implemented in software.

12. In regard to **claim 34**, the Specification (¶48) provides evidence that Applicant intends the "means for sampling ...", "means for calculating ...", "means for choosing ..." and "means for setting ..." to be implemented in software.

13. Therefore, **claims 20 – 30 and 34** are directed towards software, per se. The claims lack the necessary physical articles or objects to constitute a machine or a

manufacture within the meaning of 35 USC 101. They are clearly not a series of steps or acts to be a process nor are they a combination of chemical compounds to be a composition of matter. As such, they fail to fall within a statutory category. They are, at best, functional descriptive material *per se*. Descriptive material can be characterized as either "functional descriptive material" or "nonfunctional descriptive material." Both types of "descriptive material" are nonstatutory when claimed as descriptive material *per se*, 33 F.3d at 1360, 31 USPQ2d at 1759. When functional descriptive material is recorded on some computer-readable medium, it becomes structurally and functionally interrelated to the medium and will be statutory in most cases since use of technology permits the function of the descriptive material to be realized. Compare *In re Lowry*, 32 F.3d 1579, 1583-84, 32 USPQ2d 1031, 1035 (Fed. Cir. 1994). Merely claiming nonfunctional descriptive material, i.e., abstract ideas, stored on a computer-readable medium, in a computer, or on an electromagnetic carrier signal, does not make it statutory. See *Diehr*, 450 U.S. at 185-86, 209 USPQ at 8 (noting that the claims for an algorithm in *Benson* were unpatentable as abstract ideas because "[t]he sole practical application of the algorithm was in connection with the programming of a general purpose computer.").

Claim Rejections - 35 USC § 102(a)

14. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

15. Claims **1, 2, 5, 6, 8, 11, 13, 16, 17, 20, 24, 26, 27 and 28** are rejected under 35 U.S.C. 102(a) as being anticipated by Schleimer et al ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

16. In regard to **claim 1**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

17. In regard to **claim 2**, Schleimer et al discloses compacting the subset of the overlapping blocks includes setting bits in the representation of the document based on the subset of the overlapping blocks ((g) of Figure 2, bits of the representation of the document i.e. document signature are set).

18. In regard to **claim 5**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

19. In regard to **claim 6**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

20. In regard to **claim 8**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

21. In regard to **claim 11**, Schleimer et al discloses the overlapping blocks can be of the same length (Section 3).

22. In regard to **claim 13**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping samples (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3); and compacting the subset of the overlapping blocks includes setting bits in the representation of the document based on the subset of the overlapping blocks ((g) of Figure 2, bits of the representation of the document i.e. document signature are set).

23. In regard to **claim 16**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)) and choosing the smallest hash value in a window (Section 3).

24. In regard to **claim 17**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

25. In regard to **claim 20**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so

does Section 3). Schleimer et al further discloses comparing the document representation against a query based on the representation of another document (Section 3.2).

26. In regard to **claim 24**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

27. In regard to **claim 26**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

28. In regard to **claim 27**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

29. In regard to **claim 28**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

Claim Rejections - 35 USC § 102(b)

30. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

31. Claim **20** is rejected under 35 U.S.C. 102(b) as being anticipated by Burrows (US Patent 5,745,900 B1; patented 28 April 1998).

32. In regard to **claim 20**, Burrows discloses a fingerprint creation unit (2410) that samples the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); chooses a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows further discloses a similarity detection component to compare fingerprints to determine whether pairs of fingerprints correspond to near-duplicate documents (2420 and see figure 24).

Claim Rejections - 35 USC § 103

33. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

34. Claims **1 – 2, 5 – 8 and 10** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

35. In regard to **claim 1**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a

subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

36. In regard to **claim 2**, Burrows teaches setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

37. In regard to **claim 5**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

38. In regard to **claim 6**, Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of

ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

39. In regard to **claim 7**, Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

40. In regard to **claim 8**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

41. In regard to **claim 10**, Ward et al teaches wherein setting the bits includes flipping a bit in the representation of the document when the bit corresponds to a block in the subset of plurality of overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks).

42. Claims **3, 4, 11 and 12** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows and Ward et al as applied to claim 1 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

43. In regard to **claim 3**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

44. In regard to **claim 4**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

45. In regard to **claim 11**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a predetermined size such as 8 bytes (see col 6, lines 4 – 7). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined

length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

46. In regard to **claim 12**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a predetermined size such as 8 bytes (see col 6, lines 4 – 7) and under sized words can be padded to bring them to correct size. It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

47. Claims **13, 16 and 17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

48. In regard to **claim 13**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the

selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

49. In regard to **claim 16**, Ward et al teaches hashing of the data blocks (see ¶40). Ward et al also teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40).It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

50. In regard to **claim 17**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

51. Claims **14 and 15** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows and Ward et al as applied to claim 13 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

52. In regard to **claim 14**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the

representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

53. In regard to **claim 15**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

54. Claims **22 and 23** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows as applied to claim 20 above, and further in view of Charikar ("Similarity Estimation Techniques from Rounding Algorithms"; published 19 May 2002).

55. In regard to **claim 22**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of

fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

56. In regard to **claim 23**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

57. Claims **24 and 25** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) as applied to claim 20 above, and further in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

58. In regard to **claim 24**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be

chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

59. In regard to **claim 25**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

60. Claims **26 – 29** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

61. In regard to **claim 26**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

62. In regard to **claim 27**, Burrows teaches the invention as substantially claimed. Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

63. In regard to **claim 28**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

64. In regard to **claim 29**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

65. Claim **31, 32 and 34** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

66. In regard to **claim 31**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a

subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

67. In regard to **claim 32**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

68. In regard to **claim 34**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to

use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

69. Claim **9** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 8 above, and further in view of Official Notice.

70. In regard to **claim 9**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

71. Claim **18** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 16 above, and further in view of Official Notice.

72. In regard to **claim 18**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of

invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

73. Claim **19** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 17, and further in view of Official Notice.

In regard to **claim 19**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

74. Claim **33** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 32 above, and further in view of Official Notice.

75. In regard to **claim 33**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

76. Claim **21** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) as applied to claim 20 above, and further in view of Official Notice.

77. In regard to **claim 21**, Burrows also discloses a search engine (140, see figure 1). However, Burrows does not explicitly disclose returning a single link when the documents are determined to be duplicates. The Examiner takes Office Notice that returning a single link when the documents are determined to be duplicates is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use returning a single link with the components of Burrows because it would reduce the amount of data traffic and provide the user with clarity as to the nature of the document.

78. Claim **30** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB

2002/0133499 A1; published 19 September 2002) as applied to claim 27, and further in view of Official Notice.

79. In regard to **claim 30**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

Conclusion

80. The Examiner requests, in response to this Office action, that support be shown for language added to any original claims on amendment and any new claims. That is, indicate support for newly added claim language by specifically pointing to page(s) and line no(s) in the specification and/or drawing figure(s). This will assist the Examiner in prosecuting the application.

81. When responding to this Office action, Applicant is advised to clearly point out the patentable novelty which he or she thinks the claims present, in view of the state of the art disclosed by the references cited or the objections made. He or she must also show how the amendments avoid such references or objections See 37 CFR 1.111(c).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Garrett A. Smith whose telephone number is (571) 270-1764. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tim T. Vo can be reached on (571) 272-3642. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

February 4, 2008

Garrett Smith
Patent Examiner
Art Unit 2168

TIM VO
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100